CLAIMS

1. A method for detection the parts of the non-rigid object such as, but not limited to, human body, and for recognition gestures and activities of the object of interest in real-time video sequences comprising:

a) Eliminating the background to obtain the foreground objects;

b) Detecting different regions of the foreground objects by using color and/or shape information;

c) Finding the contours of the areas cited in 1b);

d) Fitting closed curves to the contours cited in 1c);

e) Computing unary and binary attributes for the closed curves cited in 1d);

f) Comparing the attributes cited in 1e) with the object attributes in the training data set via a matching algorithm and determining object parts after matching;

g) Combining adjacent segments and repeating claims 1c) through 1f);

h) Storing 2D center of gravity coordinates of each object part in the buffers for certain number of frames;

i) Comparing the change of center of gravity coordinates with time for each object part cited in 1h) with the templates in the training data set and recognizing the activity of each part separately;

j) Combining the activities of the object parts cited in 1h) and recognizing the overall activity of the object of interest in the scene; and,

k) Combining the gestures and activities of different objects to detect the event in the scene.

2. The method of claim 1 wherein step 1a) further comprises:

a) Grabbing and digitizing several video frames under different lighting changes;

b) Converting input color components into a single color space or a combination of color spaces such as, but not limited to, red-green-blue color space, luminance-chrominance color space, hue-saturation color space, etc., or a combination of them;

c) Generating a statistical model for background frames by using the mean and standard deviation of frame blocks and color components cited in 2b);

d) Grabbing and digitizing test video frame and repeating step 2b) for the test frame;

e) Generating a statistical model for test frame by using the mean and standard deviation of frame blocks and color components cited in 2d); and,

f) Comparing the mean and/or standard deviation of the frame blocks of the test frame cited in 2e) with the mean and/or standard deviation of the background frames cited in 2c).

3. The method of claim 2 wherein step 2f) further comprises:

eliminating the blocks with similar mean and standard deviation that are below a threshold from the test frame and generating a foreground region.

4. The method of claim 3 further comprising:

Segmenting foreground regions cited in 3 of the object of interest hierarchically into its smaller unique parts based on the combination of color components and statistical shape features.

5. The method of claim 4 wherein the segmenting step based on statistical shape features further comprises:

Comparing the curvature, mean, deviation of the object of interest with the known model shape features for different objects.

6. The method of claim 4 further comprising:

a) Assigning a number to pixels contained in the segmented foreground regions and assigning another number to the non-foreground regions and generating a binary image; and,

b) Determining and storing contour coordinates of the segmented regions cited in claims 1b) and 6a) in a buffer.

7. The method of claim 6 wherein step 6b) further comprises:

a) Initializing multiple 3 by 3 pixel windows;

b) Shifting the windows through the binary image cited in 6a) in different and independent directions; and,

c) Finding the next neighboring point on the outer boundary of the foreground object segment cited in 4a) after one of the window centers overlaps with the foreground object segment.

8. The method of claim 1 wherein step 1d) further comprises:

Approximating contours by fitting closed curves with shape preserving deformations for minimizing the effect of occlusion and local deformations.

9. The method of claim 1 wherein step 1e) further comprises:

Determining geometric invariant attributes for closed curves that are maximally discriminating between objects.

10. The method of claim 1 wherein step 1f) further comprises:

a) Generating attribute feature vectors for each closed curve where each object part and meaningful combinations represent a class;

b) Determining the class of the multi-dimensional feature vector by using quadratic Mahalanobis classifier; and,

c) Evaluating a discriminant function and checking conditional rule generation and previously matched node pairs.

11. The method of claim 1 wherein step 1g) further comprises:

a) Combining or further segmenting results of low level segmentation cited in claim 2 by using model-based segmentation;

b) Managing the segmentation process using a feedback from relational representation of the object; and,

c) Combining resulting segments produced from this initial segmentation by using a bottom-up control.

12. The method of claim 1 wherein steps 1h) and 1i) further comprise:

a) Storing the center of gravity points of the detected human body parts between motion gaps in the buffer;

b) Comparing the spatio-temporal pattern of each body part movement and the combination of several patterns with the training set pattern for different body part movements by using dynamic programming and choosing the pattern with the highest probability as the recognized activity of the body part.

13. The method of claim 7 wherein the step 1j) further comprises:

a) Generating weighted feature vectors by using recognized activities of the object parts;

b) Adding speed of object parts as a new feature to feature vector; and,

c) Classifying feature vectors by using distance classifier.

14. The method of claim 13 wherein the step 1k) further comprises:
Combining activities of different object, such as but not limited to, rigid objects and human to detect the event in the scene.

15. A method for detection human body postures and for recognition global direction of the body in compressed domain, said method comprising:

a) Eliminating background to obtain the foreground object in compressed domain;

b) Windowing the foreground object cited in 15a) and scaling the window in compressed domain by using the human body proportions;

c) Comparing AC coefficients of the scaled window cited in 15b) with the AC coefficients of different human postures in the database and recognizing posture of the human in the scene; and,

d) Comparing horizontal AC coefficient differences with the vertical AC coefficient differences and obtaining global activity of the body in compressed domain.

16. The method of claim 15 wherein step 15a) further comprises:

d) Extracting quantized DCT coefficients of the Intra frames from MPEG decoder;

e) Extracting quantized DCT coefficients from JPEG decoder; and,

f) Calculating DCT coefficients in real-time.

17. The method of claim 15 wherein step 15b) further comprises:

a) Windowing the foreground object by using human body proportions for different postures; and,

b) Scaling the window in compressed domain.

18. The method of claim 15 wherein step 15c) further comprises:

a) Computing the principal components of the energy distribution of human bodies;

b) Computing eigenvectors and eigenvalues of the covariance matrix of the human body images;

c) Calculating a set of weights based on the input image and the eigenvectors by projecting the input image onto each of the eigenvectors; and,

d) Detecting human regions by computing the distance between the mean adjusted input image and its projection onto human body space.

19. The method of claim 15 wherein step 15d) further comprises:

a) Eliminating areas without motion in compressed domain; and,

b) Comparing horizontal and vertical AC coefficient differences between frames and grouping coefficient differences regarding to the human body proportion for finding the global body part activity.

20. An apparatus for implementing the activity recognition method of claim 1, said apparatus comprising:

a) A video signal processor;

b) A general purpose computer; and,

c) A special-purpose hardware unit.

21. An apparatus for implementing the activity recognition method of claim 15, said apparatus comprising:

a) A video signal processor;

b) A general purpose computer; and,

c) A special-purpose hardware unit.

22. A method for detection the rigid object parts and for recognition activities of a rigid object in real-time video sequences comprising:

a) Eliminating the background to obtain the foreground objects;

b) Detecting different regions of the foreground objects by using color and/or shape information;

c) Finding the contours of the areas cited in 22b);

d) Fitting closed curves to the contours cited in 22c);

e) Computing unary and binary attributes for the closed curves cited in 22d);

f) Repeating claims 22a) through 22e) and comparing the object attributes with the model object attributes in the training data set via a matching algorithm and determining the object parts;

g) Combining adjacent segments and repeating claim 22f);

h) Storing 2D center of gravity coordinates of rigid object parts in the buffer for certain number of frames;

i) Comparing the change of center of gravity coordinates with time for each object part cited in 22h) with the templates in the training data set and recognizing the activity of object parts; and,

j) Combining the activities of the object parts cited in 22i) and recognizing the overall activity of the object in the scene.